

# SEQ SEARCH SUMMARY

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 18:30:43 ; Search time 3117 Seconds  
(without alignments)  
15699.161 Million cell updates/sec

Title: US-09-941-945A-1  
Perfect score: 1129  
Sequence: 1 catatgcaggcgaactcctg.....ccagtgggaatccacttcgga 1129

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*

*Searched SEQ 1&2  
in DNA databases*

28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	% Query		Length	DB	ID	Description
	No.	Score				
	1	1129	100.0	328050	1 AP005275 — GenBank	AP005275 Corynebac
	2	1129	100.0	349980	6 AX127144 EP 1108790	AX127144 Sequence
	3	1129	100.0	349980	6 AX127145 "	AX127145 Sequence
	4	564	50.0	564	6 AX120753	AX120753 Sequence
	5	564	50.0	564	6 BD162870	BD162870 Novel pol.
	6	492.6	43.6	308750	1 AP005216	AP005216 Corynebac
	7	363	32.2	1156	6 AX773541	AX773541 Sequence
	8	263	23.3	1026	6 AX120754	AX120754 Sequence
	9	263	23.3	1026	6 BD162871	BD162871 Novel pol
	10	258	22.9	341553	1 BX248355	BX248355 Corynebac
	11	133.4	11.8	874	6 AR199611	AR199611 Sequence
	12	133.4	11.8	874	6 BD069092	BD069092 Nucleic a
c	13	133.4	11.8	14240	1 AE007157	AE007157 Mycobacte
c	14	133.4	11.8	316050	1 BX248346	BX248346 Mycobacte
c	15	133.4	11.8	349606	15 BX842583	Bx842583 Mycobacte
	16	84	7.4	300425	1 AP005041	AP005041 Streptomy
	17	83.2	7.4	292100	1 SCO939121	AL939121 Streptomy
	18	80.2	7.1	4692	1 SCAJ10601	AJ010601 Streptomy
c	19	62.4	5.5	42325	1 U00015	U00015 Mycobacteri
c	20	62.4	5.5	42325	6 AR345363	AR345363 Sequence
	21	62.4	5.5	344050	1 MLEPRTN2	AL583918 Mycobacte
c	22	49.2	4.4	2000	6 AX655393	AX655393 Sequence
c	23	43	3.8	125020	9 AF429315	AF429315 Homo sapi
	24	42.8	3.8	125020	9 AF429315	AF429315 Homo sapi
	25	42	3.7	193584	2 AC124172	AC124172 Mus muscu
c	26	41.8	3.7	98466	2 AF551762	AF551762 Mus muscu
c	27	41.8	3.7	165036	10 AC126433	AC126433 Mus muscu
c	28	41.8	3.7	191713	10 AC117241	AC117241 Mus muscu
	29	41.8	3.7	198325	2 AY211543	AY211543 Mus muscu
	30	40.6	3.6	2000	6 AX655393	AX655393 Sequence
c	31	40	3.5	314450	1 BX294147	BX294147 Pirellula
	32	39.4	3.5	277000	1 SCO939109	AL939109 Streptomy
	33	38.8	3.4	5610	6 AX278007	AX278007 Sequence

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 18:25:39 ; Search time 387 Seconds  
 (without alignments)  
 12393.322 Million cell updates/sec

Title: US-09-941-945A-1  
 Perfect score: 1129  
 Sequence: 1 catatgcaggcgaactcctg.....ccagtgggaatccacttcgga 1129

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
 1: geneseqn1980s:\*  
 2: geneseqn1990s:\*  
 3: geneseqn2000s:\*  
 4: geneseqn2001as:\*  
 5: geneseqn2001bs:\*  
 6: geneseqn2002s:\*  
 7: geneseqn2003as:\*  
 8: geneseqn2003bs:\*  
 9: geneseqn2003cs:\*  
 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1129	100.0	349980	5 AAH68525	Aah68525 C glutami
2	1129	100.0	349980	5 AAH68526	Aah68526 C glutami
3	1125.8	99.7	1129	6 ABA96000	Aba96000 Corynebac
4	564	50.0	564	5 AAH65634	Aah65634 C glutami
5	564	50.0	564	7 ACA00076	Aca00076 C. glutam
6	263	23.3	1026	5 AAH65635	Aah65635 C glutami
7	263	23.3	1026	7 ACA00077	Aca00077 C. glutam
8	133.4	11.8	1105	2 AAV44589	Aav44589 Mycobacte

EP 1108790

"

DE 10043331

Applicants

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 22:19:59 ; Search time 80 Seconds  
 (without alignments)  
 7831.744 Million cell updates/sec

Title: US-09-941-945A-1  
 Perfect score: 1129  
 Sequence: 1 catatgcaggcgaactcctg.....ccagtggaatccacttcgga 1129

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

		%					Description
Result	Query	Match	Length	DB	ID		
No.	Score						
	1	133.4	11.8	874	4	US-09-082-920-1	Sequence 1, Appli
c	2	133.4	11.8	4403765	3	US-09-103-840A-2	Sequence 2, Appli
c	3	133.4	11.8	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	4	62.4	5.5	42325	4	US-08-311-731A-131	Sequence 131, App
	5	38.8	3.4	5610	4	US-10-204-708-54	Sequence 54, Appl
	6	37.2	3.3	813	4	US-09-252-991A-5906	Sequence 5906, Ap
c	7	37.2	3.3	1605	4	US-09-252-991A-5966	Sequence 5966, Ap
c	8	37.2	3.3	1623	4	US-09-252-991A-5984	Sequence 5984, Ap
c	9	36.6	3.2	684	4	US-09-252-991A-2524	Sequence 2524, Ap
	10	36.2	3.2	966	4	US-09-252-991A-2236	Sequence 2236, Ap
c	11	35.6	3.2	2016	4	US-09-252-991A-2445	Sequence 2445, Ap
	12	35.4	3.1	1429	4	US-08-879-337-2	Sequence 2, Appli

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 23:47:24 ; Search time 401 Seconds  
(without alignments)  
12844.123 Million cell updates/sec

Title: US-09-941-945A-1  
Perfect score: 1129  
Sequence: 1 catatgcaggcgaactcctg.....ccagtggaaatccacttcgga 1129

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2995936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

2002/0197605  
FD 12/2000

Result		%					
No.	Score	Query Match	Length	DB	ID	Description	
1	1129	100.0	1129	9	US-09-941-945A-1	Sequence 1, Appli	
2	1129	100.0	3309400	9	US-09-738-626-1	Sequence 1, Appli	
3	564	50.0	564	9	US-09-738-626-669	Sequence 669, App	
4	263	23.3	1026	9	US-09-738-626-670	Sequence 670, App	
5	84	7.4	9025608	15	US-10-156-761-1	Sequence 1, Appli	
6	81.6	7.2	573	15	US-10-156-761-4983	Sequence 4983, Ap	
7	38.8	3.4	5610	15	US-10-239-676-170	Sequence 170, App	
8	38.8	3.4	5610	15	US-10-204-708-54	Sequence 54, Appl	
9	38.8	3.4	5610	15	US-10-240-453-192	Sequence 192, App	
10	38.8	3.4	5610	17	US-10-240-589C-90	Sequence 90, Appl	
11	38.4	3.4	2320	13	US-10-027-632-264352	Sequence 264352,	
12	38.4	3.4	2320	16	US-10-027-632-264352	Sequence 264352,	
c 13	38.2	3.4	2655	13	US-10-282-122A-15067	Sequence 15067, A	
c 14	37.8	3.3	2562	13	US-10-282-122A-31875	Sequence 31875, A	
15	37.8	3.3	18011	15	US-10-311-455-8	Sequence 8, Appli	
16	36.6	3.2	1002	16	US-10-369-493-45411	Sequence 45411, A	
17	36.4	3.2	671	15	US-10-184-644-346	Sequence 346, App	
18	36.4	3.2	671	15	US-10-184-634-346	Sequence 346, App	
c 19	36.4	3.2	737	13	US-10-425-114-27636	Sequence 27636, A	
c 20	35.6	3.2	453	9	US-09-938-842A-3437	Sequence 3437, Ap	
c 21	35.6	3.2	453	11	US-09-938-842A-3437	Sequence 3437, Ap	
c 22	35.6	3.2	1704	13	US-10-282-122A-30661	Sequence 30661, A	
23	35.4	3.1	1429	8	US-08-879-337-1	Sequence 1, Appli	
24	35.2	3.1	1449	13	US-10-424-599-137657	Sequence 137657,	
25	35	3.1	716	13	US-10-142-426-96	Sequence 96, Appl	
26	35	3.1	716	15	US-10-123-155-96	Sequence 96, Appl	
27	35	3.1	716	15	US-10-146-731-96	Sequence 96, Appl	
28	35	3.1	716	15	US-10-140-472-96	Sequence 96, Appl	
29	35	3.1	716	15	US-10-141-761-96	Sequence 96, Appl	
30	35	3.1	716	15	US-10-142-885-96	Sequence 96, Appl	
31	35	3.1	716	15	US-10-158-790-96	Sequence 96, Appl	
32	35	3.1	716	16	US-10-137-871-96	Sequence 96, Appl	
33	35	3.1	716	16	US-10-140-923-96	Sequence 96, Appl	
34	35	3.1	716	16	US-10-141-756-96	Sequence 96, Appl	
35	35	3.1	716	16	US-10-141-759-96	Sequence 96, Appl	
36	35	3.1	716	16	US-10-140-805-96	Sequence 96, Appl	
37	35	3.1	716	16	US-10-140-864-96	Sequence 96, Appl	
c 38	35	3.1	4639	15	US-10-096-534-48	Sequence 48, Appl	
c 39	34.8	3.1	745	13	US-10-424-599-142608	Sequence 142608,	
c 40	34.8	3.1	771	13	US-10-194-163-667	Sequence 667, App	
41	34.8	3.1	1407	13	US-10-417-700A-76	Sequence 76, Appl	
42	34.6	3.1	723	16	US-10-369-493-42286	Sequence 42286, A	
c 43	34.6	3.1	1040	13	US-10-425-114-21645	Sequence 21645, A	
44	34.6	3.1	1710	16	US-10-369-493-27553	Sequence 27553, A	
c 45	34.6	3.1	2199	13	US-10-282-122A-25416	Sequence 25416, A	

## ALIGNMENTS

RESULT 1  
US-09-941-945A-1

OM nucleic - nucleic search, using sw model

Run on: June 6, 2004, 22:09:53 ; Search time 2427 Seconds  
(without alignments)  
13891.389 Million cell updates/sec

Title: US-09-941-945A-1  
Perfect score: 1129  
Sequence: 1 catatgcaggcgaactcctg.....ccagtggaatccacttcgga 1129

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
	1	48.4	4.3	1201	13	BX361080	BX361080 BX361080
c	2	45.2	4.0	1207	29	CNS015ZN	AL106109 Drosophil
	3	41	3.6	885	13	BX425603	BX425603 BX425603
	4	39.8	3.5	1201	13	BX364960	BX364960 BX364960
c	5	39.2	3.5	1201	13	BX381961	BX381961 BX381961
	6	39	3.5	712	13	BX416727	BX416727 BX416727
c	7	39	3.5	1101	29	CNS00LOO	AL068607 Drosophil
c	8	39	3.5	1101	29	CNS00ZB7	AL097453 Drosophil
	9	38.6	3.4	1101	29	CNS00LT2	AL078714 Drosophil
c	10	38.2	3.4	395	13	BY224949	BY224949 BY224949
c	11	38	3.4	1201	13	BX356004	BX356004 BX356004
	12	37.6	3.3	1201	13	BX381961	BX381961 BX381961
	13	37.6	3.3	2072	11	AK016917	AK016917 Mus muscu
	14	37.6	3.3	2243	11	AK029525	AK029525 Mus muscu
	15	37.4	3.3	866	29	CG771362	CG771362 TcB47.2_F
c	16	37.2	3.3	435	14	W96254	W96254 ze42b01.rl
c	17	37.2	3.3	444	9	AA018279	AA018279 ze53h06.r
c	18	37.2	3.3	460	9	AA018267	AA018267 ze53g07.r
c	19	37.2	3.3	464	9	AA054016	AA054016 zf48h07.r
c	20	37.2	3.3	493	9	AA013384	AA013384 ze28b12.r
c	21	37.2	3.3	505	14	H86191	H86191 ys94g06.rl
	22	37.2	3.3	807	12	BG321006	BG321006 Zm04_03e1
	23	37.2	3.3	887	12	BI096866	BI096866 SCUMtig11
	24	37.2	3.3	933	28	BZ554206	BZ554206 pacs1-60
c	25	37.2	3.3	1045	13	BX386287	BX386287 BX386287
c	26	37.2	3.3	1086	28	BZ550840	BZ550840 pacs1-60
c	27	37	3.3	389	14	H84094	H84094 ys94f11.rl
c	28	37	3.3	400	14	H84088	H84088 ys94e12.rl
c	29	37	3.3	427	9	AA056119	AA056119 zf55g07.r
c	30	37	3.3	995	13	BX404378	BX404378 BX404378
	31	36.8	3.3	1101	29	CNS00FCQ	AL070525 Drosophil
c	32	36.6	3.2	421	28	AQ234900	AQ234900 HS_2054_A
c	33	36.6	3.2	475	10	BF412036	BF412036 UI-R-BT1-
	34	36.6	3.2	531	12	BP135277	BP135277 BP135277
c	35	36.6	3.2	704	29	AG171291	AG171291 Pan trogl
	36	36.6	3.2	884	29	CNS006U0	AL065923 Drosophil
	37	36.6	3.2	947	28	BZ570083	BZ570083 msh2_1157
	38	36.4	3.2	289	10	AW372518	AW372518 PM3-BT034
c	39	36.4	3.2	541	28	AQ611746	AQ611746 HS_5087_B
c	40	36.4	3.2	565	28	B68684	B68684 CIT-HSP-205
c	41	36.2	3.2	442	14	CF485590	CF485590 POL1_32_A
c	42	36.2	3.2	552	29	CG335625	CG335625 OG3BN04TV
	43	36.2	3.2	559	29	CG335610	CG335610 OG3BN04TH
c	44	36.2	3.2	650	29	CC610253	CC610253 OGUCS45TV
c	45	36.2	3.2	706	12	BM338529	BM338529 MEST228-C



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 7, 2004, 00:31:14 ; Search time 2368 Seconds  
(without alignments)  
3441.085 Million cell updates/sec

Title: US-09-941-945A-2  
Perfect score: 929  
Sequence: 1 MADTERELADLVPQATAGDR.....VAQHRALTTLRSTLEQQENK 188

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

Q=/cgn2\_1/USPTO\_spool/US09941945/runat\_03062004\_164702\_5297/app\_query.fasta\_1.32  
7

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09941945\_@CGN\_1\_1\_2527\_@runat\_03062004\_164702\_5297 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*

11: gb\_sts:\*  
 12: gb\_sy:\*  
 13: gb\_un:\*  
 14: gb\_vi:\*  
 15: em\_ba:\*  
 16: em\_fun:\*  
 17: em\_hum:\*  
 18: em\_in:\*  
 19: em\_mu:\*  
 20: em\_om:\*  
 21: em\_or:\*  
 22: em\_ov:\*  
 23: em\_pat:\*  
 24: em\_ph:\*  
 25: em\_pl:\*  
 26: em\_ro:\*  
 27: em\_sts:\*  
 28: em\_un:\*  
 29: em\_vi:\*  
 30: em\_htg\_hum:\*  
 31: em\_htg\_inv:\*  
 32: em\_htg\_other:\*  
 33: em\_htg\_mus:\*  
 34: em\_htg\_pln:\*  
 35: em\_htg\_rod:\*  
 36: em\_htg\_mam:\*  
 37: em\_htg\_vrt:\*  
 38: em\_sy:\*  
 39: em\_htgo\_hum:\*  
 40: em\_htgo\_mus:\*  
 41: em\_htgo\_other:\*

although its "TTG" which is Leu

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the test score distribution.

calls first residue MET

# SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	926	99.7	564	6	AX120753	EP1108790 AX120753 Sequence
2	926	99.7	564	6	BD162870	JP 20021913 70 BD162870 Novel pol
3	926	99.7	328050	1	AP005275	- GenBank AP005275 Corynebac
4	926	99.7	349980	6	AX127144	EP AX127144 Sequence
5	926	99.7	349980	6	AX127145	EP AX127145 Sequence
6	850	91.5	308750	1	AP005216	AP005216 Corynebac
7	657	70.7	341553	1	BX248355	BX248355 Corynebac
8	473	50.9	874	6	AR199611	AR199611 Sequence
9	473	50.9	874	6	BD069092	BD069092 Nucleic a
c 10	473	50.9	14240	1	AE007157	AE007157 Mycobacte
c 11	473	50.9	316050	1	BX248346	BX248346 Mycobacte
c 12	473	50.9	349606	15	BX842583	Bx842583 Mycobacte
13	397.5	42.8	300425	1	AP005041	AP005041 Streptomy
14	388	41.8	292100	1	SC0939121	AL939121 Streptomy
15	385	41.4	4692	1	SCAJ10601	AJ010601 Streptomy
c 16	320	34.4	42325	1	U00015	U00015 Mycobacteri

7/2002

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 7, 2004, 00:29:29 ; Search time 301 Seconds  
(without alignments)  
2653.360 Million cell updates/sec

Title: US-09-941-945A-2  
Perfect score: 929  
Sequence: 1 MADTERELADLVPQATAGDR.....VAQHRALTTLRSTLEQQENK 188

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

Q=/cgn2\_1/USPTO\_spool/US09941945/runat\_03062004\_164701\_5287/app\_query.fasta\_1.32  
7

-DB=N\_Geneseq\_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09941945 @CGN\_1\_1\_352 @runat\_03062004\_164701\_5287 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002s:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result	No.	Score	% Query Match Length DB	ID	Description
	1	926	99.7 564 5	AAH65634	Aah65634 C glutami
	2	926	99.7 564 7	ACA00076	Aca00076 C. glutam
	3	926	99.7 349980 5	AAH68525	Aah68525 C glutami
	4	926	99.7 349980 5	AAH68526	Aah68526 C glutami
	5	920	99.0 1129 6	ABA96000	Aba96000 Corynebac
	6	473	50.9 1105 2	AAV44589	Aav44589 Mycobacte
c	7	473	50.9 110000 4	AAI99682_38	Continuation (39 o
c	8	473	50.9 110000 4	AAI99683_38	Continuation (39 o
c	9	320	34.4 42325 9	ADB74382	Adb74382 Mycobacte
	10	153.5	16.5 564 6	ABK74063	Abk74063 Bacillus
	11	145.5	15.7 639 5	AAH66206	Aah66206 C glutami
	12	145.5	15.7 648 7	ACA00415	Aca00415 C. glutam
	13	145.5	15.7 1330 6	AAD35114	Aad35114 Corynebac
	14	145.5	15.7 2086 6	AAD35115	Aad35115 Corynebac
	15	145.5	15.7 349980 5	AAH68527	Aah68527 C glutami
	16	135.5	14.6 577 6	ABK74074	Abk74074 Bacillus
	17	129	13.9 585 7	ABZ42002	Abz42002 N. gonorr
	18	127	13.7 110000 4	AAI99682_08	Continuation (9 of
	19	127	13.7 110000 4	AAI99683_08	Continuation (9 of
	20	123	13.2 37096 3	AAA81489_8	Continuation (9 of
	21	123	13.2 92934 3	AAA81473	Aaa81473 N. mening
	22	123	13.2 172325 3	AAF21613	Aaf21613 Neisseria
	23	122	13.1 13336 4	AAS59554	Aas59554 Propionib
	24	122	13.1 13336 7	ACF64483	Acf64483 Propionib
c	25	121	13.0 14615 4	AAS59577	Aas59577 Propionib
c	26	121	13.0 14615 7	ACF64506	Acf64506 Propionib
	27	121	13.0 37856 3	AAA11992	Aaa11992 S. cellul
	28	119.5	12.9 38675 9	ADB74386	Adb74386 Mycobacte
	29	116.5	12.5 647 2	AAQ70249	Aaq70249 P. aerugi
	30	116.5	12.5 647 2	AAQ70242	Aaq70242 P. aerugi
	31	116.5	12.5 647 3	AAA51919	Aaa51919 P. aerugi
	32	116.5	12.5 771 6	ABT05614	Abt05614 Mycobacte
	33	116.5	12.5 771 7	ABZ71111	Abz71111 Mycobacte
	34	116.5	12.5 75216 6	ABX09141	Abx09141 Mycobacte
	35	116.5	12.5 110000 4	AAI99682_13	Continuation (14 o
	36	116.5	12.5 110000 4	AAI99683_13	Continuation (14 o
	37	116	12.5 603 3	AAA64868	Aaa64868 Bordetell
c	38	116	12.5 35026 3	AAA64890	Aaa64890 Bordetell
	39	111	11.9 717 8	ADA30919	Ada30919 DNA encod
	40	111	11.9 110000 2	AAT42063_06	Continuation (7 of
	41	109	11.7 135638 7	ABX34289	Abx34289 S. atrool
	42	108.5	11.7 582 7	AAD47841	Aad47841 Haemophil
	43	108.5	11.7 603 5	AAH68381	Aah68381 C glutami
	44	108.5	11.7 603 7	ACA01297	Aca01297 C. glutam
	45	108.5	11.7 1211 6	AAL42355	Aal42355 Corynebac

# ALIGNMENTS

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 7, 2004, 01:25:39 ; Search time 66 Seconds  
(without alignments)  
1580.769 Million cell updates/sec

Title: US-09-941-945A-2  
Perfect score: 929  
Sequence: 1 MADTERELADLVPQATAGDR.....VAQHRALTTLRSTLEQQENK 188

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

Q=/cgn2\_1/USPTO\_spool/US09941945/runat\_03062004\_164703\_5323/app\_query.fasta\_1.32  
7

-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09941945\_@CGN\_1\_1\_54\_@runat\_03062004\_164703\_5323 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result			%		Query		DB	ID	Description
	No.	Score	Match	Length					
	1	473	50.9	874	4	US-09-082-920-1			Sequence 1, Appli
c	2	473	50.9	4403765	3	US-09-103-840A-2			Sequence 2, Appli
c	3	473	50.9	4411529	3	US-09-103-840A-1			Sequence 1, Appli
c	4	320	34.4	42325	4	US-08-311-731A-131			Sequence 131, App
	5	139	15.0	621	4	US-09-252-991A-1369			Sequence 1369, Ap
	6	127	13.7	4403765	3	US-09-103-840A-2			Sequence 2, Appli
	7	127	13.7	4411529	3	US-09-103-840A-1			Sequence 1, Appli
	8	121	13.0	597	4	US-09-252-991A-13522			Sequence 13522, A
	9	119.5	12.9	38675	4	US-08-311-731A-135			Sequence 135, App
	10	116.5	12.5	647	1	US-08-260-202A-9			Sequence 9, Appli
	11	116.5	12.5	647	1	US-08-017-114-9			Sequence 9, Appli
	12	116.5	12.5	647	3	US-08-505-307-9			Sequence 9, Appli
	13	116.5	12.5	647	4	US-09-609-151A-9			Sequence 9, Appli
	14	116.5	12.5	647	5	PCT-US94-02034-9			Sequence 9, Appli
	15	116.5	12.5	840	4	US-09-252-991A-13069			Sequence 13069, A
	16	113	12.2	600	4	US-09-543-681A-3510			Sequence 3510, Ap
	17	111	11.9	717	4	US-09-328-352-2206			Sequence 2206, Ap
	18	111	11.9	1830121	4	US-09-557-884-1			Sequence 1, Appli
	19	111	11.9	1830121	4	US-09-643-990A-1			Sequence 1, Appli
c	20	108.5	11.7	1830121	4	US-09-557-884-1			Sequence 1, Appli
c	21	108.5	11.7	1830121	4	US-09-643-990A-1			Sequence 1, Appli
c	22	104	11.2	1746	4	US-09-252-991A-8985			Sequence 8985, Ap
	23	103.5	11.1	4605	4	US-09-221-017B-128			Sequence 128, App
	24	99.5	10.7	549	4	US-09-252-991A-11688			Sequence 11688, A
	25	99	10.7	846	4	US-09-252-991A-9635			Sequence 9635, Ap
c	26	99	10.7	1062	4	US-09-252-991A-9532			Sequence 9532, Ap
	27	99	10.7	1371	4	US-09-252-991A-9717			Sequence 9717, Ap
	28	97.5	10.5	540	4	US-09-252-991A-15549			Sequence 15549, A
	29	96.5	10.4	1230025	4	US-09-198-452A-1			Sequence 1, Appli
	30	96	10.3	525	4	US-09-107-532A-1956			Sequence 1956, Ap
	31	95	10.2	1173	4	US-09-107-532A-3567			Sequence 3567, Ap
	32	94	10.1	2000	1	US-08-622-353-1			Sequence 1, Appli
	33	94	10.1	2000	1	US-08-622-352A-1			Sequence 1, Appli
	34	94	10.1	2000	3	US-08-826-390-1			Sequence 1, Appli
	35	93.5	10.1	501	4	US-09-252-991A-9107			Sequence 9107, Ap
	36	92	9.9	615	4	US-09-489-039A-4559			Sequence 4559, Ap
	37	88.5	9.5	804	4	US-09-489-039A-3407			Sequence 3407, Ap
c	38	88.5	9.5	1092	4	US-09-252-991A-9447			Sequence 9447, Ap
	39	88.5	9.5	1149	4	US-09-252-991A-3630			Sequence 3630, Ap
	40	88.5	9.5	1287	4	US-09-252-991A-9384			Sequence 9384, Ap
c	41	88.5	9.5	1785	4	US-09-252-991A-9417			Sequence 9417, Ap
	42	88.5	9.5	1842	4	US-09-252-991A-9328			Sequence 9328, Ap
	43	88	9.5	80161	3	US-09-036-987A-1			Sequence 1, Appli
	44	88	9.5	80161	3	US-09-370-700-1			Sequence 1, Appli
	45	88	9.5	80161	4	US-09-603-207-1			Sequence 1, Appli

## ALIGNMENTS

RESULT 1

US-09-082-920-1

; Sequence 1, Application US/09082920C

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 7, 2004, 02:26:35 ; Search time 300 Seconds  
(without alignments)  
2858.851 Million cell updates/sec

Title: US-09-941-945A-2  
Perfect score: 929  
Sequence: 1 MADTERELADLVPQATAGDR.....VAQHRALTTLRSTLEQQENK 188

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2995936 seqs, 2280998010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

Q=/cgn2\_1/USPTO\_spool/US09941945/runat\_03062004\_164704\_5412/app\_query.fasta\_1.32  
7

-DB=Published\_Applications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09941945\_@CGN\_1\_1\_354\_@runat\_03062004\_164704\_5412  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published\_Applications\_NA:\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq2:\*  
 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result			%		Query		DB	ID	Description
	No.	Score	Match	Length	Length	Match			
	1	926	99.7	564	9	US-09-738-626-669			Sequence 669, App
	2	926	99.7	1129	9	US-09-941-945A-1			Sequence 1, Appli
	3	926	99.7	3309400	9	US-09-738-626-1			Sequence 1, Appli
	4	397.5	42.8	573	15	US-10-156-761-4983			Sequence 4983, Ap
	5	397.5	42.8	9025608	15	US-10-156-761-1			Sequence 1, Appli
	6	210.5	22.7	774	15	US-10-156-761-4719			Sequence 4719, Ap
c	7	210.5	22.7	9025608	15	US-10-156-761-1			Sequence 1, Appli
	8	210	22.6	711	15	US-10-156-761-3372			Sequence 3372, Ap
	9	194.5	20.9	687	15	US-10-156-761-3374			Sequence 3374, Ap
	10	153.5	16.5	564	9	US-09-974-300-1354			Sequence 1354, Ap
	11	145.5	15.7	537	15	US-10-156-761-4690			Sequence 4690, Ap
	12	145.5	15.7	639	9	US-09-738-626-1241			Sequence 1241, Ap
	13	145.5	15.7	1330	9	US-09-935-757-1			Sequence 1, Appli
	14	145.5	15.7	2086	9	US-09-935-757-5			Sequence 5, Appli
	15	145	15.6	612	15	US-10-156-761-886			Sequence 886, App
	16	142	15.3	651	15	US-10-156-761-606			Sequence 606, App
	17	141	15.2	618	15	US-10-156-761-6054			Sequence 6054, Ap
	18	135.5	14.6	577	9	US-09-974-300-1365			Sequence 1365, Ap
	19	135	14.5	615	15	US-10-156-761-521			Sequence 521, App
	20	121.5	13.1	657	15	US-10-156-761-3104			Sequence 3104, Ap
	21	118	12.7	537	15	US-10-156-761-5382			Sequence 5382, Ap
	22	118	12.7	618	15	US-10-156-761-6317			Sequence 6317, Ap
	23	117.5	12.6	537	15	US-10-156-761-4048			Sequence 4048, Ap
	24	116.5	12.5	75216	15	US-10-080-170-646			Sequence 646, App
	25	116	12.5	546	15	US-10-156-761-6559			Sequence 6559, Ap
	26	113	12.2	594	15	US-10-156-761-5107			Sequence 5107, Ap
	27	112.5	12.1	645	15	US-10-156-761-691			Sequence 691, App
	28	111	11.9	1830121	15	US-10-329-960-1			Sequence 1, Appli
	29	111	11.9	1830121	16	US-10-329-670-1			Sequence 1, Appli
	30	109	11.7	135638	15	US-10-314-657-1			Sequence 1, Appli
	31	108.5	11.7	603	9	US-09-738-626-3416			Sequence 3416, Ap
	32	108.5	11.7	1211	9	US-09-942-935-1			Sequence 1, Appli
c	33	108.5	11.7	1830121	15	US-10-329-960-1			Sequence 1, Appli
c	34	108.5	11.7	1830121	16	US-10-329-670-1			Sequence 1, Appli
	35	107.5	11.6	579	9	US-09-738-626-284			Sequence 284, App
	36	107.5	11.6	1002	15	US-10-156-761-3350			Sequence 3350, Ap
	37	107.5	11.6	1109	9	US-09-941-936A-1			Sequence 1, Appli
c	38	107.5	11.6	3309400	9	US-09-738-626-1			Sequence 1, Appli
	39	107	11.5	591	15	US-10-156-761-6314			Sequence 6314, Ap

USPAP  
2002/0197605



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: June 7, 2004, 01:22:34 ; Search time 2348 Seconds  
(without alignments)  
2391.009 Million cell updates/sec

Title: US-09-941-945A-2  
Perfect score: 929  
Sequence: 1 MADTERELADLVPQATAGDR.....VAQHRALTTLRSTLEQQENK 188

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

Q=/cgn2\_1/USPTO\_spool/US09941945/runat\_03062004\_164702\_5307/app\_query.fasta\_1.32  
7

-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09941945\_CGN\_1\_1\_2607@runat\_03062004\_164702\_5307 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*

```

12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result			%				ID	Description
	No.	Score	Query Match	Length	DB			
c	1	120	12.9	879	11	BC022210	BC022210	Homo sapi
	2	110	11.8	1012	28	BZ573065	BZ573065	msh2_2928
	3	109.5	11.8	763	13	BQ968973	BQ968973	QHB35M11.
	4	108.5	11.7	698	13	BQ916398	BQ916398	QHB17M04.
	5	108.5	11.7	728	13	BQ916763	BQ916763	QHB19B03.
	6	108.5	11.7	762	13	BQ969036	BQ969036	QHB35P11.
	7	107.5	11.6	701	13	BQ914474	BQ914474	QHB10K23.
	8	107.5	11.6	762	13	BU027838	BU027838	QHG8B12.y
	9	107.5	11.6	769	13	BQ916644	BQ916644	QHB18I24.
	10	105.5	11.4	739	13	BQ916439	BQ916439	QHB17N22.
c	11	103.5	11.1	717	13	BQ970657	BQ970657	QHB42L15.
	12	102.5	11.0	1115	28	BZ553599	BZ553599	pacs1-60_
	13	101.5	10.9	743	13	BQ915857	BQ915857	QHB15P24.
	14	101.5	10.9	893	13	BQ931778	BQ931778	AGENCOURT
	15	100.5	10.8	758	13	BQ917182	BQ917182	QHB20I24.
	16	100.5	10.8	992	28	BZ552520	BZ552520	pacs1-60_
	17	99	10.7	498	13	BQ272709	BQ272709	sao20d03.
	18	95.5	10.3	761	28	BZ558190	BZ558190	pacs1-60_
	19	94.5	10.2	816	28	BZ559495	BZ559495	pacs2-164
	20	94	10.1	634	9	AV934201	AV934201	AV934201
c	21	93	10.0	648	12	BJ471207	BJ471207	BJ471207
	22	92.5	10.0	483	14	CA548840	CA548840	C0810G09-
	23	92.5	10.0	753	13	BU023549	BU023549	QHF11L03.
	24	91.5	9.8	229	28	AZ578624	AZ578624	25f08 Sho
	25	91.5	9.8	767	10	BF862931	BF862931	963039F08
	26	90.5	9.7	652	13	BQ915718	BQ915718	QHB15I21.
	27	90	9.7	1001	29	CG348903	CG348903	OGWGL80TV
	28	89	9.6	2636	11	AK002623	AK002623	Mus muscu
	29	88.5	9.5	501	14	CB972586	CB972586	CAB30001